PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04N 7/14 — A1

(11) International Publication Number: WO 95/13680

(43) International Publication Date: 18 May 1995 (18.05.95)

(21) International Application Number: PCT/GB94/02445

(22) International Filing Date: 8 November 1994 (08.11.94)

(30) Priority Data:
9323223.9 10 November 1993 (10.11.93) GB

(71) Applicant *(for all designated States except US)*: VISUAL TECHNOLOGIES LTD. [IE/IE]; 7 Catherine Street, Waterford (IE).

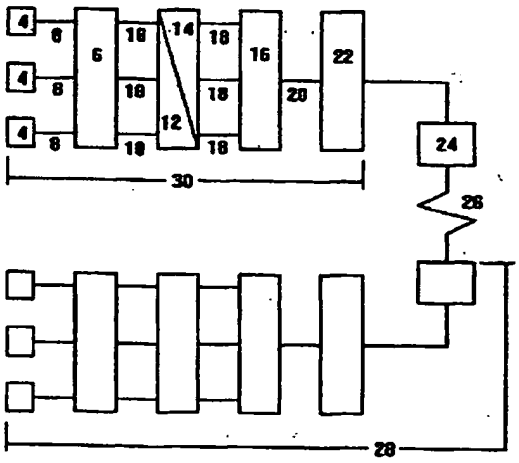
(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: RUSH, Jerome, Brian [CA/GB]; 3 Stanhope Place, London W2 2HB (GB).

(74) Agent: JONES, Graham, Henry; Graham Jones & Company, 77 Beaconsfield Road, Blackheath, London SE3 7LG (GB).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.

(54) Title: VOICE AND IMAGE TELECOMMUNICATIONS APPARATUS



(57) Abstract

Voice and image telecommunications apparatus (2) which is connectable to an analogue telephone network and which enables at least first and second persons with first and second visual display units at first and second remote locations to have a speech conversation while viewing and being able to manipulate in an interactive manner an image which is simultaneously seen by the first person at the first location and by the second person at the second location. The apparatus (2) has first connection means for the first person, second connection means for the second person, image data compression and decompression means (14), speech digitizing means (6), speech data compression and decompression means (12), combining means (16) for combining the compressed image data and the compressed speech data into a single integrated data stream, convertor means (22) for converting the single integrated data stream into analogue form suitable for transmission over the analogue telephone network, transmit and receive means (28, 30) for the first person; transmit and receive means for the second person; and control means for distributing authority over the apparatus (2) in order to ensure that only one person at a time has the ability to manipulate the dynamically changing shared image on all visual display units in operation.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

VOICE AND IMAGE TELECOMMUNICATIONS APPARATUS

This invention relates to voice and image telecommunications apparatus.

Visual display units are well known and they are widely used for displaying information such for example as words and/or images on a screen. An image can be displayed on the screen of different visual display units at remote locations but problems arise when persons operating these different visual display units wish to discuss and alter the displayed image. More specifically, known telecommunications apparatus does not allow for the simultaneous alteration in real-time of the shared image between persons in the same free and easy manner as the image may be discussed by the persons.

It is an aim of the present invention to obviate or reduce the above mentioned problem.

Accordingly, in one non-limiting embodiment of the invention there is provided voice and image telecommunications apparatus which is connectible to an analogue telephone network and which enables at least first and second persons with first and second visual display units at first and second remote locations to have a speech conversation whilst viewing and being able to manipulate in an interactive manner an image which is simultaneously seen by the first person at the first location and by the second person at the second location, which apparatus comprises:

i) first connection means which is for the first person and which is for providing a connection to the analogue telephone network and to a first visual display unit to be used by the first person;

ii) second connection means which is for the second person and which is for providing a connection to the analogue telephone network and to a second visual display unit to be used by the second person;

iii) image data compression and de-compression means for compressing and de-compressing image data;

iv) speech digitising means for converting speech into digitised speech data;

v) speech data compression and de-compression means for compressing and de-compressing the speech data;

vi) combining means for combining the compressed image data and the compressed speech data into a single integrated data stream;

vii) convertor means for converting the single integrated stream into analogue form suitable for transmission over the analogue telephone network;

viii) transmit and receive means for the first person;

ix) transmit and receive means for the second person; and

x) control means for distributing authority over the voice and image telecommunications apparatus in order to ensure that only one person at a time has the ability to manipulate the dynamically changing shared image on

all visual display units in operation during operation of the voice and image telecommunications apparatus.

It will be appreciated that the apparatus of the present invention enables the image to be altered in an interactive manner by all persons party to an operative group studying the image. The image can be altered at will by different persons substantially in the same free and easy manner as they can discuss the image at will, with the proviso that whilst all the persons in the party can alter the image, only one person at a time can do so in order to prevent two or more persons altering the image at any one time and thereby causing confusion.

Preferably, the apparatus is one in which the manipulation of the image is possible in the following two ways:

a) the original data can be altered which then changes the displayed image, or;

b) lines and markings can be drawn as an overlay to highlight parts of the image which has no effect on the original data.

Preferably, the apparatus is one in which the control means is such as to allow compressed speech data signals from the first person and from the second person to pass simultaneously in both directions in order to permit simultaneous speech between the first and the second persons, in which the control means is such as to allow compressed image data signals from the first person and from the second person to pass one at a time in both directions in order to permit only one person at a time to have the ability to manipulate the image; and in which the control means is such as to permit the transfer of the ability to manipulate the image from person to person. Preferably, the apparatus is one in which there are four data streams consisting of digitised and compressed voice data, processed and compressed

image or bitmap data, vector data which communicates the position of the cursor on the screen, and systems or control data which conveys instructions to the receiving screen to draw a line or rectangle for example, which are integrated into one combined data stream.

The apparatus may be one in which the first and the second transmit and receive means are each such as to send only an image as seen at the moment of display on the visual display unit, and in which the apparatus is such that the first and the second transmit and receive means receive what has been transmitted but with the control means giving the person transmitting control over a screen of the visual display unit of the person receiving and with the control means permitting the transfer of the ability to transmit from person to person. In this case, the first and the second transmit and receive means may each include a read only memory for sending instructions, a random access memory for extracting visual display unit screen data, and at least one microprocessor. Where two microprocessors are employed, then one microprocessor may be for speech data compression and the other microprocessor may be for speech and image data integration.

In an alternative embodiment of the invention, the first and the second transmit and receive means are each such as to transmit the interactive manipulation of transmitted data.

The first and the second transmit and receive means may be such that they each comprise a separate transmitter and a separate receiver, or they may alternatively each comprise a transceiver.

Advantageously, the image data compressing means is a hierarchical image data compressing means which compresses image data in a hierarchical order of importance for allowing transmission to a visual display unit in a similar hierarchical order of importance thereby to enable the build up of an image in a layered manner. This enables the provision of an image on the second visual display unit which is on the first visual display unit and which has not lost quality. The hierarchical transmission enables first the outline of the image to be formed and then the colour information to be inserted as a result of the coarse detail of the image being transmitted first and then the finer detail information thus providing the successive build up of the image. The apparatus may thus be one in which the image data compressing means is a hierarchical image data compressing means which compresses the image data in a shape orientated build process which obtains the approximate outline block colour representation of the image by assessing object size and shape. The technique employed may include shape and text recognition, shape repeat, filters and line approximations. The outline plus fill colour block colour representation may often fully describe the image but, where it does not, further build up techniques may be applied which include major gradients, texture simulation and repeat, secondary outlines and fill.

The hierarchical image data compressing means may provide a window of finite time duration for the total bandwidth to be devoted to the transmission of image data to allow at least sufficient image data to be transmitted for a recognisable image to be displayed by the receiving display unit, before speech interrupts. The window of finite time duration may be of any suitable and appropriate time duration such for example as 3.5 seconds.

The image data may be transferred in the following hierarchical stages:

(i) image outline;

(ii) fill colour;

(iii) refined colour;

(iv) refined image; and

(v) repeat stages (iii) and (iv) until the required image is obtained.

The image data is preferably transferred in the following hierarchical stages:

(i) colour plates providing coarse shape approximations

(ii) smaller shapes derived from a codebook

(iii) refinement of image detail from colour plates of narrower colour value variations

(iv) refinement of the image using major gradients, texture simulation and repeat, secondary outlines and fill techniques

(v) repeat stages (iii) and (iv) until the required image is obtained.

The apparatus of the invention may include dynamic partitioning means for producing data packets in order to allow the integration of the data in a dynamic manner to provide for the efficient transmission of a data stream over the available bandwidth.

The apparatus with the dynamic partitioning means may also include data packet shrinking means for shrinking the data packets as the line quality changes or drops in quality or, data packet size changing means for changing the size of the data packets as the line quality changes in order to minimise the impact of all data lost since lost visual data can be re-transmitted but lost speech data is permanently lost.

The apparatus may further include dynamic data packet sizing means for dynamically sizing the data packets as line quality changes in order to minimise on the input of data that is either lost or that require re-transmission, and in order to maximise on the transmission of data. Such apparatus enables the optimum transmission of data.

The apparatus of the invention may have the following four data streams which have to be integrated;

a) voice data - digitised and compressed

b) image or bitmap data (compressed)

c) vector data - the position of the cursor on the screen

d) systems or control data - instructions to the receiving screen to draw a line or rectangle for example.

Usually, streams b, c and d are first combined into a single stream and than a difference processor combines the 'combined' data stream and the voice data into a single data stream. It is necessary to inform the receiving computer of the nature and amount of the data being transmitted and that information is contained in bytes in data packet 'headers' and 'enders'.

It is desirable to decrease the percentage of overhead as a proportion of the data being transmitted to as low a level as possible. The overhead is always a fixed amount and so in conditions of good line quality, the data packet would be as long as possible to minimise that percentage. As quality drops however, there will be a requirement to minimise the damage of lost data. Voice is lost forever but image, vector and control data is re-transmitted. It is better to re-transmit small amounts

more often than large amounts because of the time delay and synchronisation problems. It is thus possible to adjust the data packet size.

The apparatus of the invention may include monitor means for monitoring for a gap in the transmitted speech for the insertion of extra image data.

An embodiment of the invention will now be described solely by way of example and with reference to accompanying drawings in which:

Figure 1 is a block diagram of voice and image telecommunications apparatus;

Figure 2 is a block diagram of a printed circuit board for installation into a visual display unit for converting the visual display unit into voice and image telecommunications apparatus; and

Figure 3 is a block circuit diagram similar to that shown in Figure 2.

Referring to Figure 1, there is shown voice and image telecommunications apparatus 2 which is connectible to an analogue telephone network and which enables at least first and second persons with first and second visual display units at first and second remote locations to have a speech conversation whilst viewing and being able to manipulate in an interactive manner an image which is seen by the first person at the first location and by the second person at the second location. The apparatus 2 comprises data capture means 4 for capturing information from a telephone line and/or the visual display units.

The data capture means 4 is connected to speech digitising means 6 via input lines 8. The speech digitising means 6 is for converting speech into digitised speech data. The speech digitising means 6 is connected by input lines 10 to speech data

compression means 12 and also to image data compression means 14. The dual speech compression means 12 and the image data compression means 14 are connected to combining means 16 by lines 18. The combining means 16 combines the compressed image data and the compressed speech data into a single integrated data stream which passes along line 20 to convertor means 22. The convertor means 22 is for converting the single integrated data stream passing along line 20 into analogue form suitable for transmission over the analogue telephone network. The analogue form of the data may then be pumped by data pumping means 24 along line 26 to receiver means 28. It will thus be appreciated that apparatus 4, 6, 12, 14, 16, 22 forms transmitter means 30 when data is being sent to the receiver means 28.

The apparatus 2 is such that the receiver means 28 can also operate as transmitter means, in which case the transmitter means 30 will operate as receiver means 30. This is achieved by control means (not shown) for distributing authority over the apparatus 2 in order to ensure that only one person at a time has the ability to manipulate the image an all visual display units in operation during operation of the apparatus 2. More specifically, the control means may be such as to allow compressed speech data signals from the first person and from the second person to pass simultaneously in both directions in order to permit simultaneous speech between the first and the second persons. The control is independent of the transmission of speech and image data. That is the first person may be transmitting speech and image data to the second person but the second person may be exercising authority over the screen of the first person at the same time. The control means may also be such as to allow combined data stream (containing speech, image and systems data) to pass simultaneously in both directions. It is only the vector data, which is controlled by systems commands, which can be sent in only

one direction at a time. The control means will also be such as to permit the transfer of the ability to manipulate the image from person to person so that all persons party to an audio/visual discussion have the facility of making an audio and a visual contribution to the discussion. Usually, an image sent will only be the image as seen by other persons on their visual display units.

The apparatus of the present invention, for example as illustrated generally with reference to Figure 1 can be provided as a first version which combines voice and screen data to enable each user to share the same image and manipulate that image in an interactive manner with a voice contribution being possible.

In an alternative embodiment of the invention, the apparatus of the invention may be such that it does not include the voice transmission. The apparatus of the invention will then allow the interactive manipulation of the image on the visual display units but the required voice discussion will be over normal analogue telephone lines.

In the first above mentioned embodiment of the invention where the apparatus of the invention provides both voice and image discussion facilities, then the apparatus may be comprise apparatus as shown in Figure 2. Figure 2 shows a printed circuit board 32 comprising a microprocessor 34, a speech data/processor interface 36, a voice Digital Signal Processor compression device 38, a random access memory 40, a digital to analogue converter 42, an analogue to digital converter 44, and an automatic gain control 46. Items 42, 44 and 46 have been combined into a single codec which performs these three functions. The single codec is connected to a telephone handset via a digital path 50 to a telecommunication socket 54. A modem 56 is connected to communications software 58 which is in turn connected to a Digital Access Arrangement (DAA) line protector 60, which is in turn connected to

the telecommunications socket 54. A read only memory 62 and a random access memory64 are connected as shown to the microprocessor 34.

Referring now to Figure 3, similar parts as in Figure 2 have been given the same reference numerals for ease of comparison and understanding. Figure 3 illustrates how data compression of screen data can be in a host computer or in the printed circuit board 33 at position D2.

The apparatus of the present invention is able to operate such that digitised audio signals are packaged into packets and integrated with picture, drawing, numerical or textual data, layouts and other computer held or generated information, as well as images that have been scanned in or captured by image grabbing devices into a single data stream, multiplexed and distributed at the receiving end. All changes made to the screen data on either visual display unit are echoed on the screen of the other visual display unit in real time. Whilst talking, authority over both screens can be negotiated between the parties, thus enabling a document to be discussed and edited between the users as a collaborative effort, each user hearing and seeing the actions of the other.

The apparatus of the invention may use software permitting the capture of a screen of data from a non-windows application which is then converted into a bitmapped image format which can then be imported into windows software.

The apparatus of the present invention may provide users with graphic tools to assist the manipulation and annotation of screen data. All added data overlays the original and can be removed in layers. The data added by each party is preferably in their own unique colour providing an audit trail of which party has added what

annotation. Screens can at any time be saved by either party for later printing or display by either party.

Where it is desired to offer the apparatus of the invention without the voice facility, such as when for example, connection is to a local area network then the telephone handset, and the voice conversion and processing circuitry and the high bit rate modem will not be required. The apparatus will then be such as to hook up and to communicate over a local area network for the transmission and interactive manipulation of screen data and the distribution of the voice will be over the existing telephone network.

During operation of the apparatus 2 the data capture means 4 are used to generate a stream of data. Two devices may be used, namely a telephone handset and a computer. It will be appreciated that the 'telephone handset' could be any microphone that generates an analogue signal and any speaker that produces sound from an analogue signal such as a telephone handset, and the 'computer' could be extended to 'any device that digitally stores, processes and displays data such as a personal computer.

The computer will be the computer from which image data is captured. The image data is obtained from the video memory and input data is obtained from the video memory and/or keyboard buffer which can contain keyboard inputs and/or mouse inputs or inputs from other pointing devices. The screen image can be modified and the changes captured from the video memory or the keyboard buffer as they occur. The screen image can be a display of an image file or can be the screen display of an active application. In both modes, the sending and the receiving screens are slaved and authority over both screens can be traded back and forth using a system

or and agreed protocol, thus providing the users with interactive simultaneous voice and visual communications. Images can be modified or overlain with drawn annotations, and applications can be operated and up-dated in a co-operative and interactive manner in real time. Screen data can be bitmap or alpha-numeric in character.

In a possible modification of the apparatus, more than two devices may be inputting information. The apparatus may include a moving or still camera with digital or analogue input, a scanner or other electronic devices generating a data stream.

The data captured by the data capture means 4 passes along the lines 8 as shown in Figure 1. Each data capture means 4 generates a data stream. Two or more data capture means 4 may be employed.

The speech digitising means 6 may be a microphone which transmits the speech signal to the printed circuit board board as an analogue signal. The signal may be first augmented by the automatic gain control 46 to stabilise its energy level and a signal may then be converted to a digital stream using the analogue to digital converter chip 44 . In a possible future modification of the invention, the input signal may not require conversion.

The conversion process at the sending end is reversed at the receiving end by passing the digital speech stream through the speech compression and de-compression module 43 and then through the digital to analogue converter chip 42. Conversion of the voice signal is not required in the non-speech version. The data captured from the video memory is a digital data stream and does not require conversion.

The data compression is effected such that the input signal or signals are reduced in size so that when they are combined into a single data stream they do not exceed the bandwidth capacity of the transport medium. For example, a first data stream may be a digitised speech data stream which is sent to the speech compression module which consists of a digital signal processor and supporting memory chips. The data stream is compressed from 64,000 bits per second to good telephone quality produced at a rate of 6,800 bits per second or to 4,800 bits per second depending on the bandwidth that the transmission medium is able to support. The quality is that known as good telephone quality, that is good enough to be easily understood and to be able to be recognise who is speaking.

A second data stream may be data captured from the video memory which is processed and compressed in a hierarchical manner so that the critical data is extracted and sufficient data is sent to a receiving screen within a window of finite duration, for example 3.5 seconds. At the receiving screen, the whole of the bandwidth is reserved for the transmission of the screen data. During the finite window, an image is constructed on the receiving screen that accurately emulates the image on the sending screen, this image being sufficient to enable a meaningful voice/visual conversation to take place. After the expiration of the time duration of the window, the voice channel re-opens for the transmission of a combined voice and image data stream allowing voice communication and the continuing improvement and refinement of the transmitted image.

The same approach is used if the apparatus of the invention is used without the speech facility but, because the available bandwidth would be greater because there would not be the need to share it with another (voice) data stream. The transmission

of a meaningful image would take less time and the interactive manipulation of shared data could commence in less that the finite time of the window, for example the above mentioned 3.5 second window. Where there is not a voice element to the apparatus of the invention, the end of the finite window would allow for vector and systems commands to be integrated with image data and sent to the receiving screen.

With regard to the data stream known as data stream n, then depending upon the nature of the data being captured and the bandwidth both available and required, the data may be appropriately processed and compressed.

The speech data compressing means 12 and the image data compression means 14 may be such that data from data stream one and data from data stream n are packaged into packets of n bits. The data packet is flagged both front and back with information indicating that :

a) what follows will be a speech or image, vector and/or control data packet of n length and ,

b) what went before was a speech or image, vector and/or control data packet of n length.

Packet sizes are dynamically adjusted depending on the line quality as characterised by the data rejection rates. As the flag overhead is a fixed amount, it is therefore a higher proportion of a small packet than it is of a big packet. The bigger the packet, the more information data is lost in the event of an transmission fault or error but the more data can be sent in a given time period. A balance point is desired between the speed of transmission and the amount of data lost in the event of an error.

Where the apparatus of the invention offers both the speech and the image facility, the combined data may be transmitted to a similarly equipped terminal over the existing public service telephone network in analogue format. The nominal bandwidth is 9,600 bits per second but the bandwidth may range from 2,400 bits per second to 28,800 bits per second. V.32bis modem communications has become sufficiently stable that if a connection can support 2,400bps in the older V.22 standard, it can sustain 9,600bps using the V.32bis standard. We will also be compatible with the new V.FAST standard enabling data to be transferred at a rate of 28,800bps. Our effective bandwidth ranges then from 9,600 to 28,800bps. The whole of the available bandwidth may be dedicated to the transmission of the screen image for the first finite period after which the bandwidth is shared by both the voice and the image data. After compression, the voice data occupies 6,800 (or 5,200) bits per second and the screen image, vector and systems data and the data packetising overhead occupies the balance. The combined data stream is dynamically partitioned in that, in periods of voice silence, the lack of a voice signal is detected and the bandwidth no longer occupied with speech is given over to the transmission of screen data if required. When speech data is detected, the bandwidth is again segregated between image and speech data reserving 6,800 (or 5,200) bits per second for speech.

The data pumping means 24 is effective to pump the combined single stream of data to the other visual display unit over the analogue telephone network. The data must therefore be converted to an analogue signal for transmission by an analogue data transmission pump which may take the form of a modem. Modified versions of the apparatus of the invention may enable the transmission of digital data over the integrated service digital network and may therefore not require conversion.

When the apparatus of the invention is connected to a local area network, as described earlier, the data does not require conversion and the speech data conversion means 43 and the data pumping means 24 is removed.

Conversion to an external medium would not be required by the apparatus shown in the drawings.

As indicated above, the appropriate data stream is pumped by the data pumping means 24 over the analogue or digital network telephone network. The transmission conventions conform to international CCITT standards thus enabling the printed circuit boards to also function as a standard file transfer modem responding to Hayes compatible AT commands and as a Group 3, Class I and II facsimile device. If data is pumped over the digital (ISDN) network, the apparatus of the present invention will support Group 4 facsimile standards.

In the modified version of the invention where speech is not provided for, then a data pump as herein before described would not be required.

It is to be appreciated that the embodiment of the invention described with reference to the accompanying drawings has been given by way of example only and that modifications may be effected. Thus, for example, the apparatus of the invention may be used by any suitable and appropriate number of persons.

CLAIMS

1. Voice and image communications apparatus which is connectible to a telephone network and which enables at least first and second persons with first and second visual display units at first and second remote locations to have a speech conversation whilst viewing and being able to manipulate in an interactive manner the same or a shared image which is seen by the first person at the first location and by the second person at the second location, which apparatus comprises:

(i) first connection means which is for the first person and which is for providing a connection to the analogue telephone network and to a first visual display unit to be used by the first person;

(ii) second connection means which is for the second person and which is for providing a connection to the analogue telephone network and to a second visual display unit to be used by the second person;

(iii) image data compression and de-compression means for compressing and de-compressing image data;

(iv) speech digitising means for converting speech into digitised speech data;

(v) speech data compression and de-compression means for compressing and de-compressing the speech data;

(vi) combining means for combining the compressed image data and the compressed speech data into a single integrated data stream;

(vii) convertor means for converting the single integrated stream into analogue form suitable for transmission over the analogue telephone network;

(viii) transmit and receive means for the first person;

(ix) transmit and receive means for the second person; and

(x) control means for distributing authority over the voice and image telecommunications apparatus in order to ensure that only one person at a time has the ability to manipulate the dynamically changing shared image on all visual display units in operation during operation of the voice and image telecommunications apparatus.

2. Apparatus according to claim 1 in which the control means is such as to allow compressed speech data signals from the first person and from the second person to pass simultaneously in both directions in order to permit simultaneous speech between the first and the second persons, in which the control means is such as to allow compressed image data signals from the first person and from the second person to pass one at a time in both directions in order to permit only one person at a time to have the ability to manipulate the image; and in which the control means is such as to permit the transfer of the ability to manipulate the image from person to person.

3. Apparatus according to claim 2 in which there are four data streams consisting of digitised and compressed voice data, processed and compressed image or bitmap data, vector data which communicates the position of the cursor on

the screen, and systems or control data which conveys instructions to the receiving screen to draw a line or rectangle for example, which are integrated into one combined data stream.

4. Apparatus according to any one of the preceding claims in which the first and the second transmit and receive means are each such as to send only an image as seen at the moment of display on the visual display unit, and in which the apparatus is such that the first and the second transmit and receive means receive what has been transmitted but with the control means giving the person transmitting control over a screen of the visual display unit of the person receiving and with the control means permitting the transfer of the ability to transmit from person to person.

5. Apparatus according to claim 4 in which the first and the second transmit and receive means each include a read only memory for sending instructions, a random access memory for extracting visual display unit screen data, a microprocessor for speech data compression and a microprocessor for speech and image data integration.

6. Apparatus according to any one of claims 1 to 3 in which the first and the second transmit and receive means are each such as to transmit the interactive manipulation of transmitted data.

7. Apparatus according to any one of the preceding claims in which the image data compressing means is a hierarchical image data compressing means which compresses the image data in a hierarchical manner in order of importance for allowing transmission to a visual display unit in a similar hierarchical order of importance thereby to enable the build up of an image in a layered manner.

8. Apparatus according to claim 7 in which the hierarchical image data compressing means provides a window of finite time duration to allow at least sufficient image data to be transmitted for a recognisable image to be displayed by the receiving visual display unit, before speech interrupts.

9. Apparatus according to claim 7 or claim 8 in which the image data is processed, transmitted and re-constituted on the receiving visual display unit in a shape orientated build up process which obtains the approximate block colour representation of the image by assessing object size and shape, the technique employed including shape and text recognition, shape codebook, shape repeat, filters and line approximations in the following hierarchical stages:

a) a coarse detail approximation of the image made up of plates of equal block colour corresponding to large ascertained shapes;
b) finer shapes and characters determined from a variable codebook;
c) finer detail refinement consisting of areas of with decreased variations in colour values, and;
d) a successive build up of the image by repeating c) with reduced colour variations with each pass, and where this technique does not result in a fully described image, further build up techniques may include major gradients, texture simulation and repeat, secondary outlines and fill.

10. Apparatus according to any one of the preceding claims and including dynamic partitioning means for producing data packets in order to allow the

integration of the data in a dynamic manner to provide for the efficient transmission of a data stream over an available bandwidth.

11. Apparatus according to claim 10 and including data packet shrinking means for shrinking the data packet as the line quality changes or drops in quality, in order to minimise the impact of all data lost since lost visual data can be re-transmitted but lost speech data is permanently lost.

12. Apparatus according to claim 11 and including dynamic data packet sizing means for dynamically sizing the data packets as line quality changes in order to minimise on the input of data that is either lost or that requires re-transmission, and in order to maximise on the transmission of data.

13. Apparatus according to any one of claims 10 to 12 and including monitor means for monitoring for a gap in the transmitted speech for the insertion of extra image data.

14. Apparatus according to any one of the preceding claims and which is for more than two persons.

15. Voice and image communications apparatus substantially as herein described with reference to the accompanying drawings.

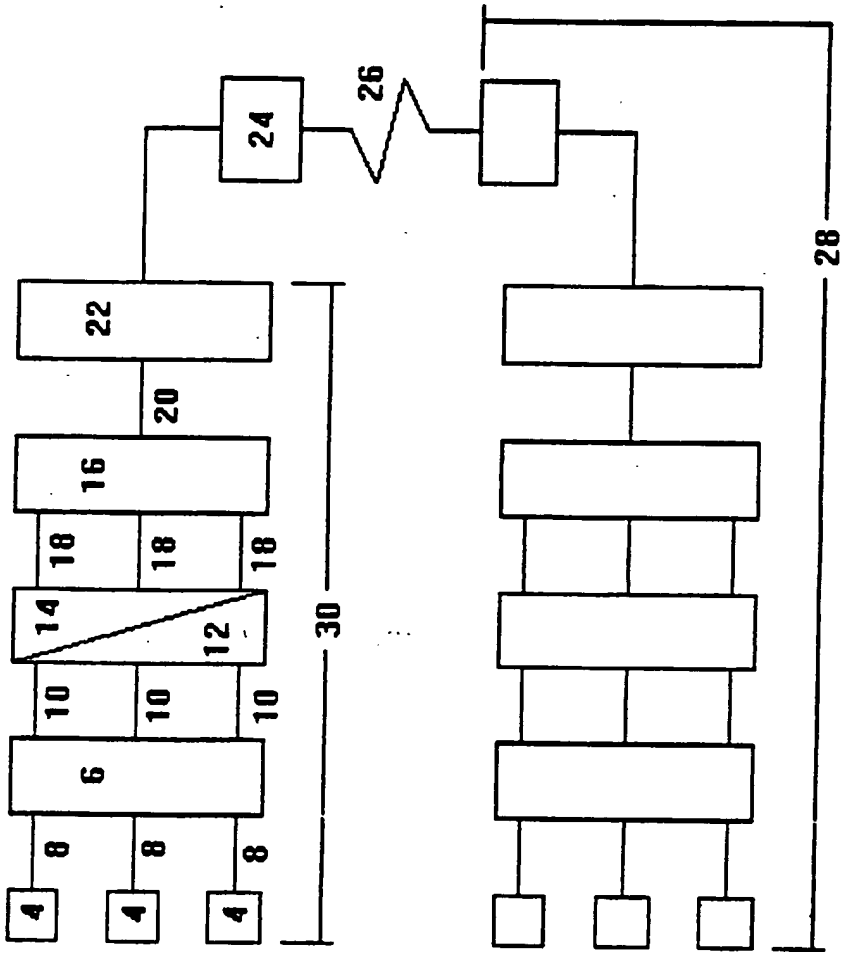


FIGURE 1

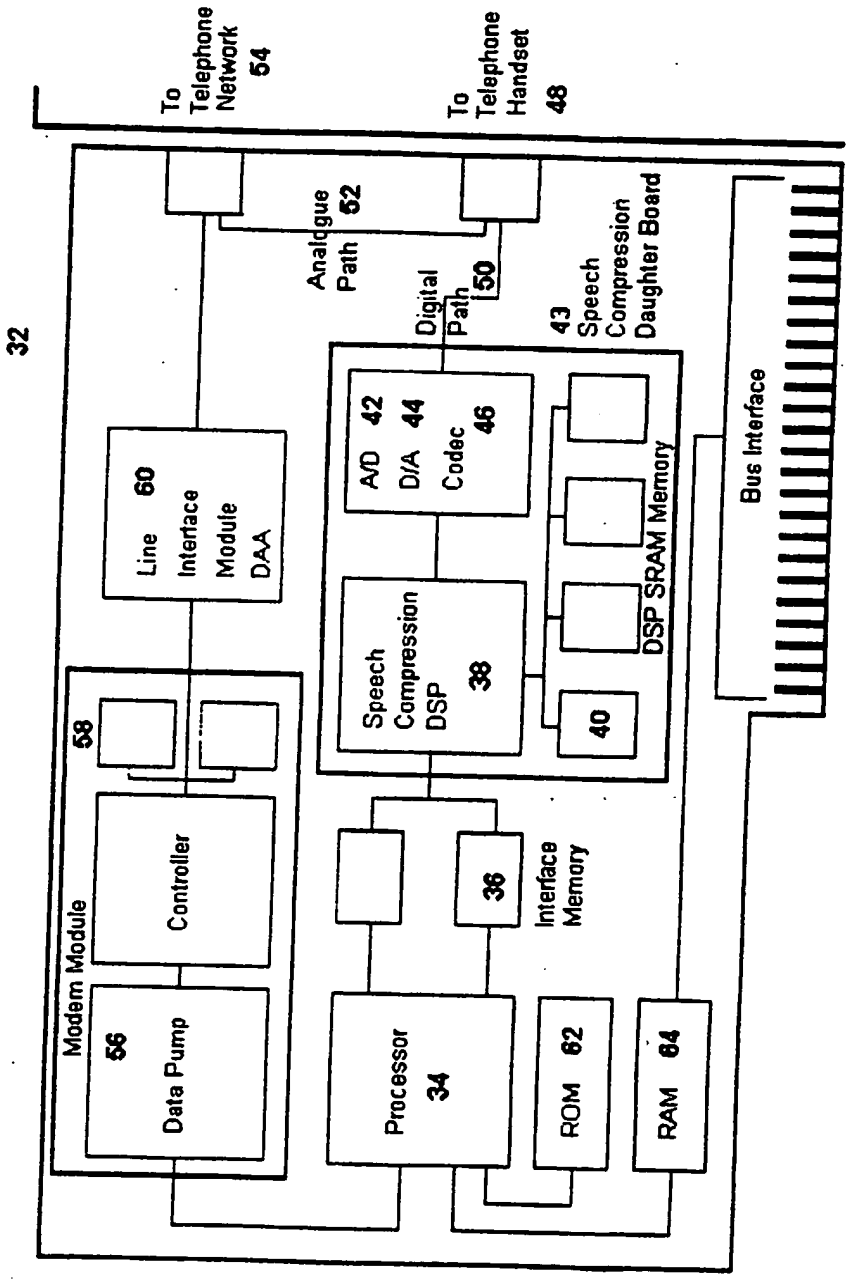

32
To Telephone Network 54
To Telephone Handset 48
Analogue Path 52
Digital Path 50
43 Speech Compression Daughter Board
Bus Interface
Line Interface Module DAA 60
A/D 42
D/A 44
Codec 46
Speech Compression DSP 38
DSP SRAM Memory
40
58
Modem Module
Controller
56 Data Pump
36
Interface Memory
Processor 34
ROM 62
RAM 64

FIGURE 2

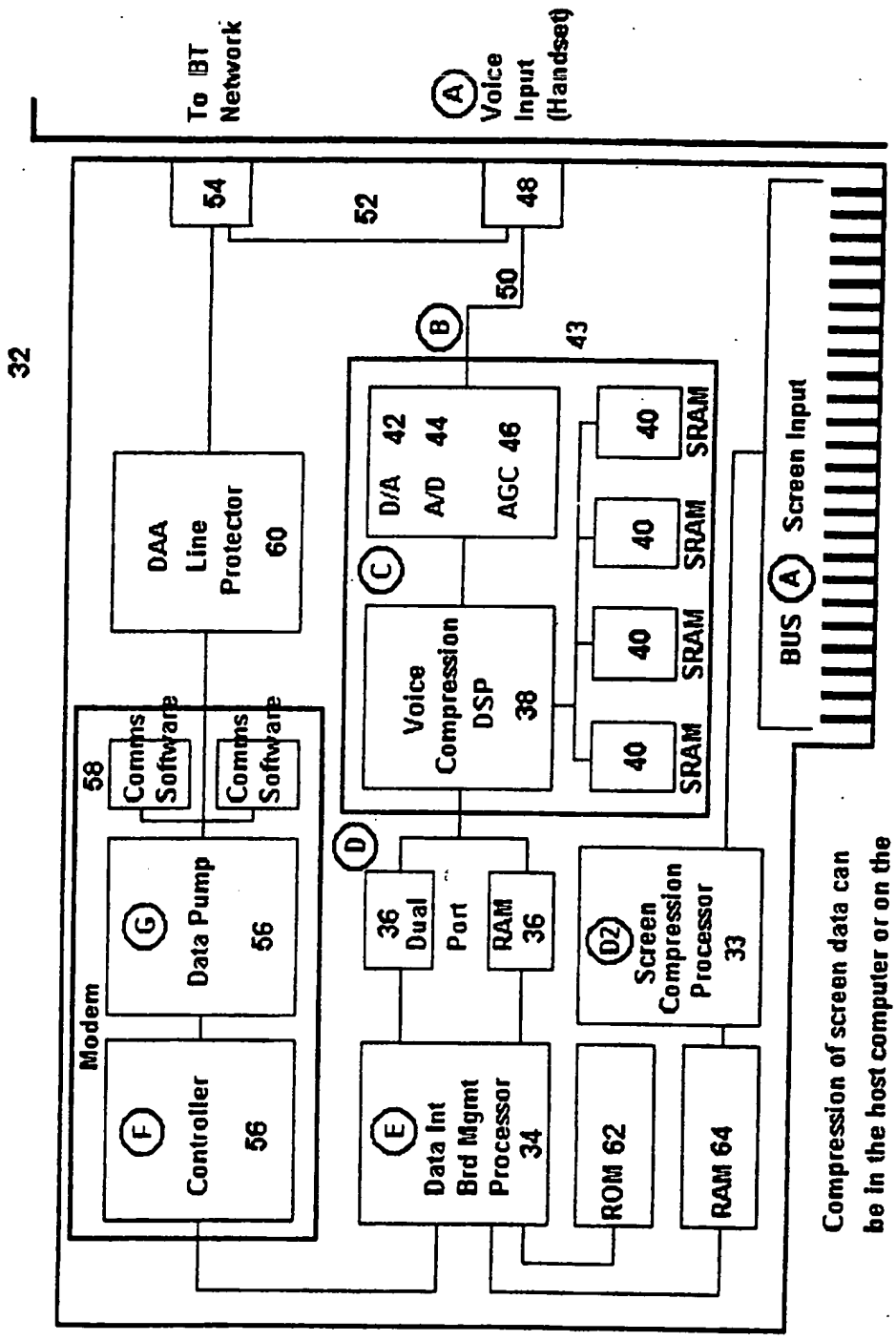

To BT Network
A Voice Input (Handset)
54
52
48
50
B
43
32
DAA Line Protector 60
D/A 42
A/D 44
AGC 46
C
Voice Compression DSP 38
40
SRAM
BUS A Screen Input
58
Comms Software
Comms Software
Modem
G Data Pump 56
F Controller 56
D
36 Dual Port
RAM 36
D2 Screen Compression Processor 33
E Data Int Brd Mgmt Processor 34
ROM 62
RAM 64
Compression of screen data can be in the host computer or on the PCB at position D2

FIGURE 3

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/GB 94/02445

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04N7/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,93 18607 (GEC-MARCONI LTD.) 16 September 1993 see page 1, line 4 - page 29, line 2; figures 1-10	1-15
A	EP,A,0 535 890 (CANON) 7 April 1993 see abstract see page 4, column 5, line 17 - page 9, column 15, line 1 see figures 1-14	1-15
A	EP,A,0 390 170 (MITSUBISHI) 3 October 1990	

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 February 1995	17. 02. 95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Van der Zaal, R

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

.nformation on patent family members

Inte mal Application No
PCT/GB 94/02445

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9318607	16-09-93	WO-A- 9318619	16-09-93
EP-A-535890	07-04-93	JP-A- 5095548	16-04-93
		US-A- 5381412	10-01-95
EP-A-390170	03-10-90	JP-A- 2257783	18-10-90
		US-A- 5063587	05-11-91